

Amendment To The Claims

Claims 1-11 Canceled.

12. (Original) A method for assessing similarity between two data objects, comprising the steps of:

- a. training a first predictive model with a first set of data objects of type X and matched data objects of type Y;
- b. using said first predictive model to assess compatibility between each of a plurality of X,Y pairs, wherein for each X,Y pair, each X is a member of a second set of data objects of type X and each Y is a member of a second set of data objects of type Y;
- c. assigning an X,Y compatibility score to each X,Y pair;
- d. comparing the X,Y compatibility scores of each member of the second set of data objects of type X with each other member of the second set of data objects of type X;
- e. pairing each member of the second set of data objects of type X with selected other members of the second set of data objects of type X having similar X,Y compatibility scores to identify a first plurality of X,X pairs, said first plurality of X,X pairs being matched pairs for training a second predictive model;
- f. selecting other ones of the second set of data objects of type X that do not have as similar compatibility scores as the matched pairs to identify a second plurality of X,X pairs, said second plurality of X,X pairs being distracters for training said second predictive model;
- g. deriving a respective set of variables from each member of the second set of data objects of type X;
- h. comparing the respective set of variables derived from each X,X matched pair and from each X,X distracter pair to determine a set of X,X comparisons;
- i. training a second predictive model with said set of X,X comparisons;

- j. receiving two data objects of type X that are not in either the first training dataset or second training dataset;
- k. deriving respective variables from each of said two data objects of type X;
- l. comparing the respective variables derived from each of said two data objects of type X to determine a production X,X comparison; and
- m. running said production X,X comparison through said second predictive model to calculate a similarity score for said two data objects of type X.

13. (Original) The method of claim 12, wherein either of said first and second predictive models comprise a respective neural network.

14. (Original) The method of claim 12, wherein either of said first and second predictive models comprise a respective regression model.

15. (Original) The method of claim 12, wherein the two data objects of type X include documents.

16. (Original) The method of claim 12, wherein the two data objects of type X include one of resumes and job descriptions.

17. (Original) The method of claim 16, wherein the respective derived variables and the respective sets of derived variables include one or more of following:

- i. reduced representation of the words in a resume;
- ii. reduced representation of the words in the education section of a resume;

- iii. reduced representation of each job description in a resume;
- iv. years of experience;
- v. standardized variables;
- vi. such as job titles;
- vii. industry SIC codes; and
- viii. degree names.

18. (Original) The method of claim 12, further comprising the steps of:
- n. repeating steps j) through m) for a plurality of production data objects of type X not in the first and second training datasets; and
 - o. clustering the plurality of production data objects according to the calculated similarity scores.
19. (Original) The method of claim 12, further comprising the steps of:
- n. repeating steps j) through m) for a plurality of production data objects of type X in a database; and
 - o. organizing the database of production data objects of type X based on the calculated similarity scores.
20. (Original) The method of claim 12, further comprising the steps of:
- n. repeating steps j) through m) for a plurality of production data objects of type X; and
 - o. deriving from the calculated similarity scores one of a measure of supply of data objects of type X and a measure of demand for a particular one of the plurality of the production data objects of type X.

21. (Original) The method of claim 12, wherein the step of comparing the respective variables further includes the steps of:

- i) constructing a vector for each of the two data objects from the derived respective variables; and
- ii) calculating the cosine of the angle between the vectors.

22. (Original) The method of claim 12, wherein the step of comparing the respective variables further includes the steps of:

- i) constructing a vector for each of the two data objects from the derived respective variables; and
- ii) calculating the dot product of the vectors.

Claims 23-36 Canceled.